



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:13 AM GMT

PDB ID : 2PMZ  
Title : Archaeal RNA polymerase from Sulfolobus solfataricus  
Authors : Murakami, K.S.  
Deposited on : 2007-04-23  
Resolution : 3.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

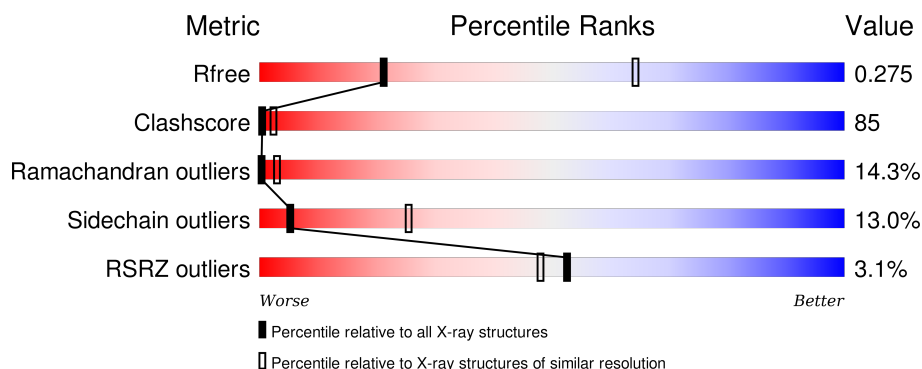
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



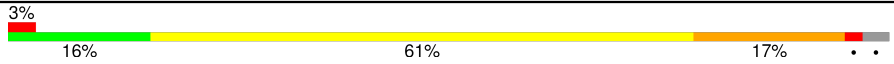
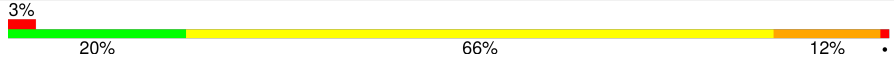
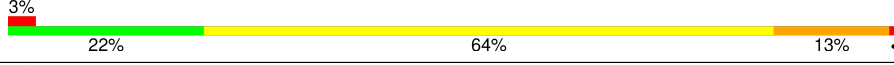
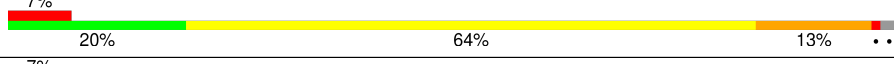
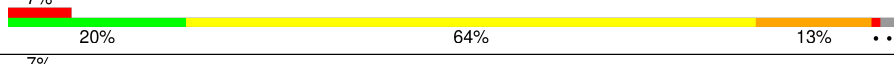
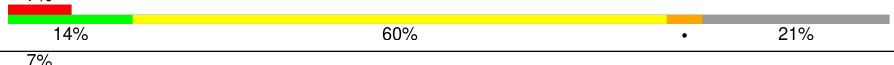
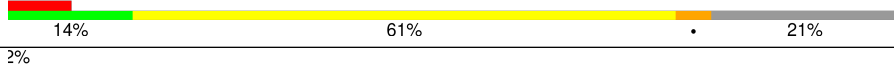
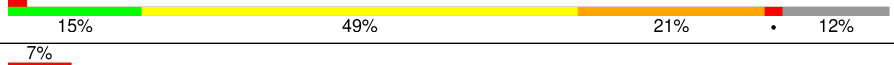
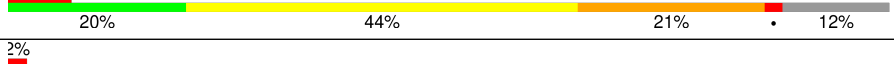
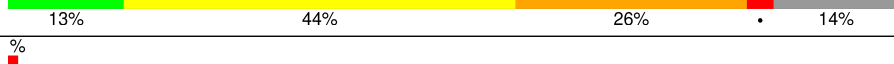

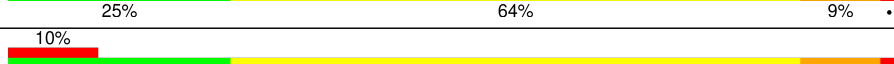
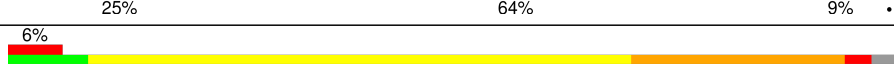
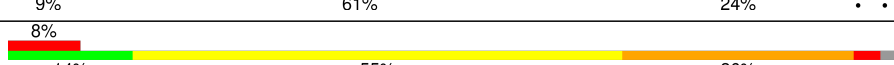
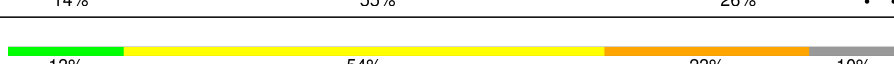
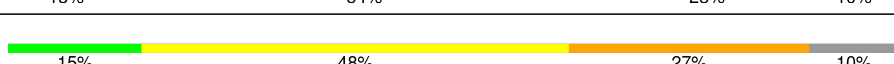
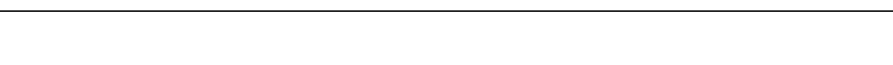
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	880	<div> <div>2%</div> <div>16% 52% 18% 12%</div> </div>
1	Q	880	<div> <div>3%</div> <div>16% 54% 17% 12%</div> </div>
2	C	392	<div> <div>2%</div> <div>11% 42% 15% 29%</div> </div>
2	G	392	<div> <div>2%</div> <div>11% 41% 16% 29%</div> </div>
3	B	1124	<div> <div>2%</div> <div>17% 60% 18% 5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	R	1124	
4	D	265	
4	S	265	
5	E	180	
5	T	180	
6	F	113	
6	U	113	
7	H	84	
7	V	84	
8	K	95	
8	W	95	
9	L	92	
9	X	92	
10	N	66	
10	Y	66	
11	P	48	
11	Z	48	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
14	F3S	D	1001	-	-	X	-
14	F3S	S	1001	-	-	X	-

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 48122 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	776	Total	C	N	O	S	0	0	0
			6173	3936	1081	1135	21			
1	Q	776	Total	C	N	O	S	0	0	0
			6173	3936	1081	1135	21			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit A".

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	279	Total	C	N	O	S	0	0	0
			2169	1376	375	412	6			
2	G	279	Total	C	N	O	S	0	0	0
			2169	1376	375	412	6			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	1090	Total	C	N	O	S	0	0	0
			8645	5483	1529	1602	31			
3	R	1090	Total	C	N	O	S	0	0	0
			8645	5483	1529	1602	31			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	264	Total	C	N	O	S	0	0	0
			2114	1355	342	403	14			
4	S	264	Total	C	N	O	S	0	0	0
			2114	1355	342	403	14			

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	176	Total	C	N	O	S	0	0	0
			1402	903	236	259	4			
5	T	176	Total	C	N	O	S	0	0	0
			1402	903	236	259	4			

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	89	Total	C	N	O	S	0	0	0
			694	433	115	142	4			
6	U	89	Total	C	N	O	S	0	0	0
			694	433	115	142	4			

- Molecule 7 is a protein called DNA-directed RNA polymerase subunit H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	74	Total	C	N	O		0	0	0
			611	397	109	105				
7	V	74	Total	C	N	O		0	0	0
			611	397	109	105				

- Molecule 8 is a protein called DNA-directed RNA polymerase subunit K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	K	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			
8	W	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			

- Molecule 9 is a protein called DNA-directed RNA polymerase subunit L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	L	92	Total	C	N	O	S	0	0	0
			723	459	121	141	2			
9	X	92	Total	C	N	O	S	0	0	0
			723	459	121	141	2			

- Molecule 10 is a protein called DNA-directed RNA polymerase subunit N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	N	64	Total	C	N	O	S	0	0	0
			514	326	94	88	6			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	Y	64	Total	C	N	O	S	0	0	0
			514	326	94	88	6			

- Molecule 11 is a protein called DNA-directed RNA polymerase subunit P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	P	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			
11	Z	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			

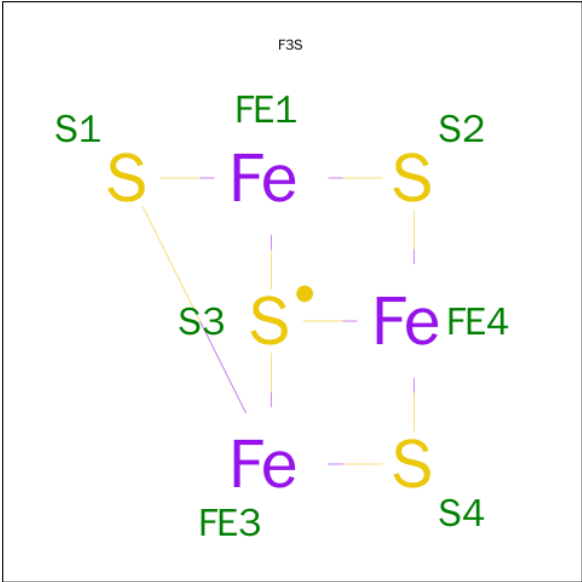
- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	P	1	Total	Zn	0	0
			1	1		
12	Q	1	Total	Zn	0	0
			1	1		
12	B	1	Total	Zn	0	0
			1	1		
12	Z	1	Total	Zn	0	0
			1	1		
12	A	1	Total	Zn	0	0
			1	1		
12	N	1	Total	Zn	0	0
			1	1		
12	R	1	Total	Zn	0	0
			1	1		
12	Y	1	Total	Zn	0	0
			1	1		

- Molecule 13 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	Q	1	Total	Mg	0	0
			1	1		
13	A	1	Total	Mg	0	0
			1	1		

- Molecule 14 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe<sub>3</sub>S<sub>4</sub>).

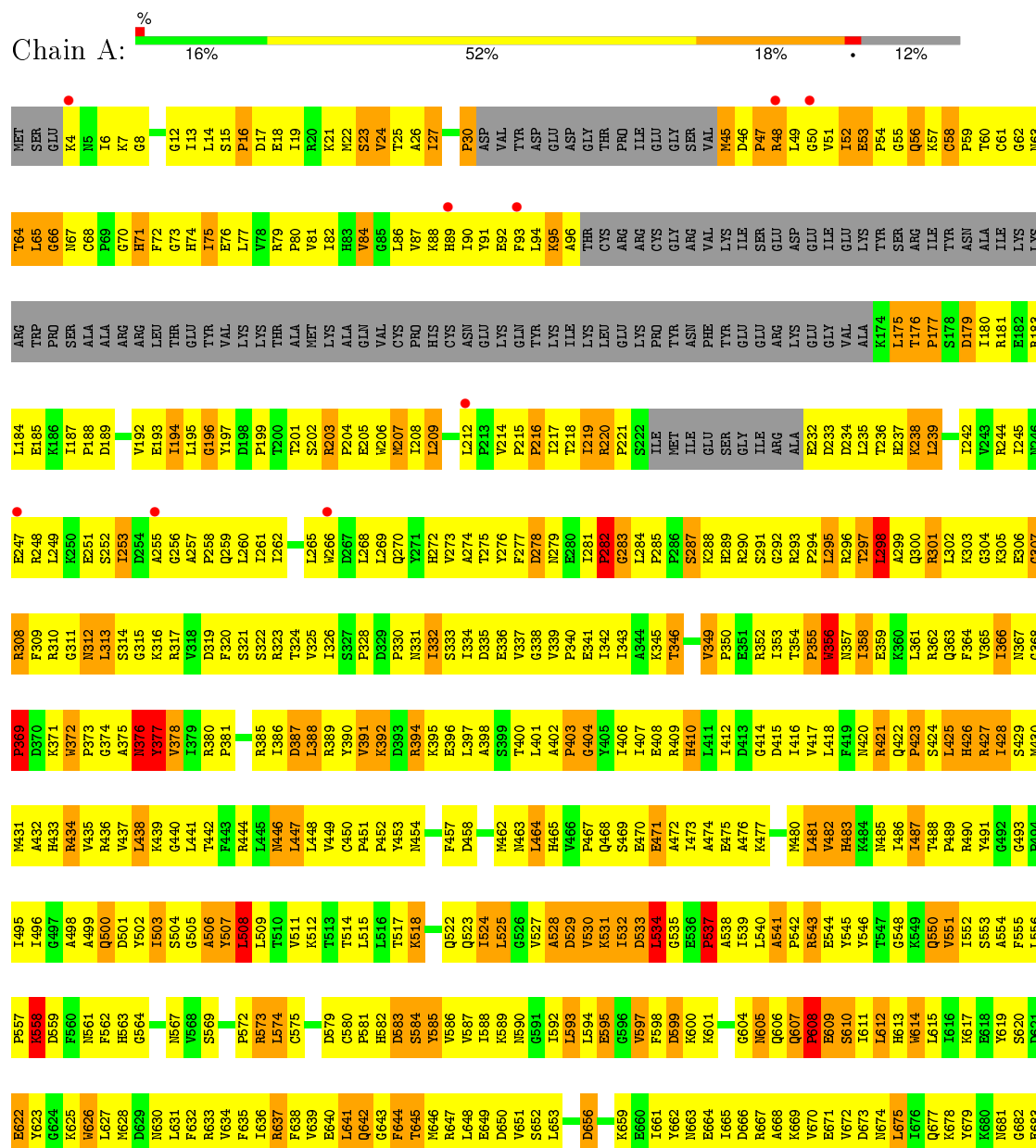


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
14	D	1	Total	Fe	S	0	0
			7	3	4		
14	S	1	Total	Fe	S	0	0
			7	3	4		

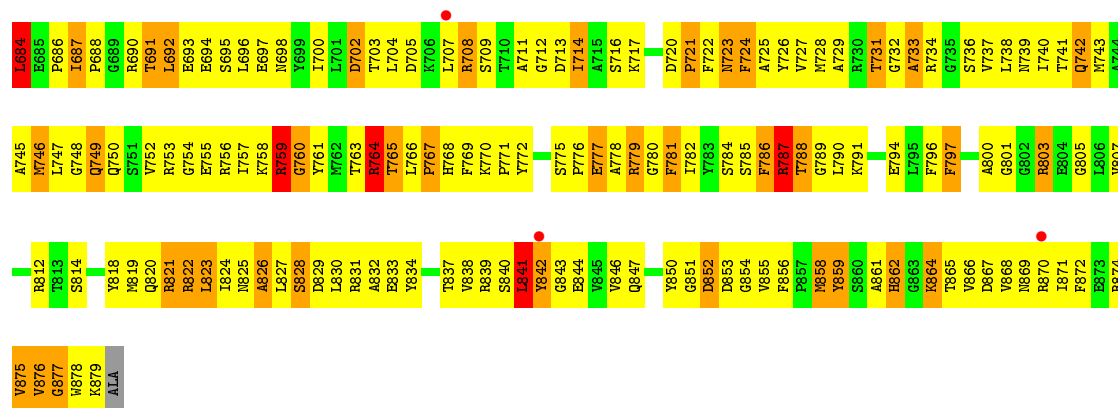
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

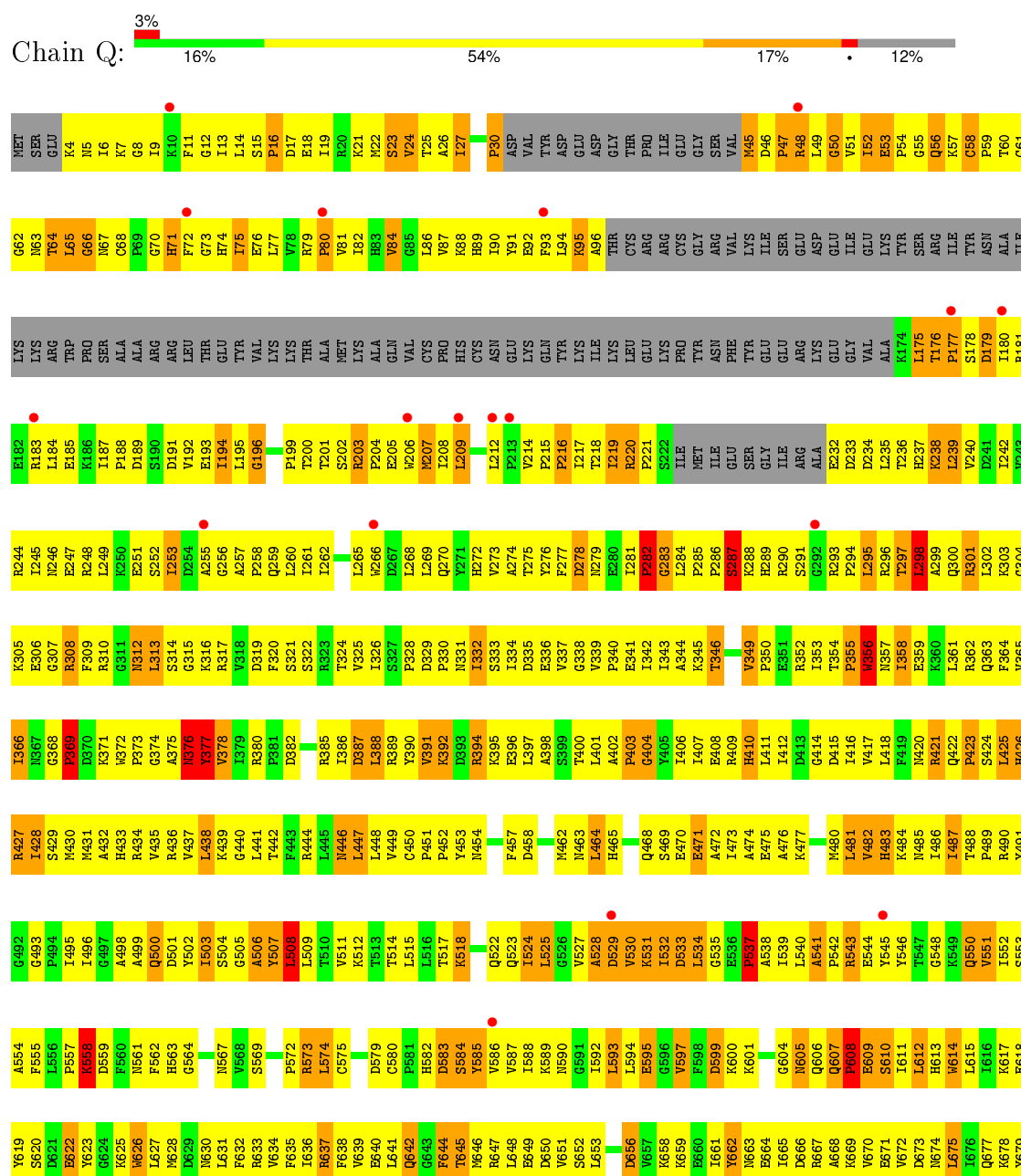
#### • Molecule 1: DNA-directed RNA polymerase subunit A

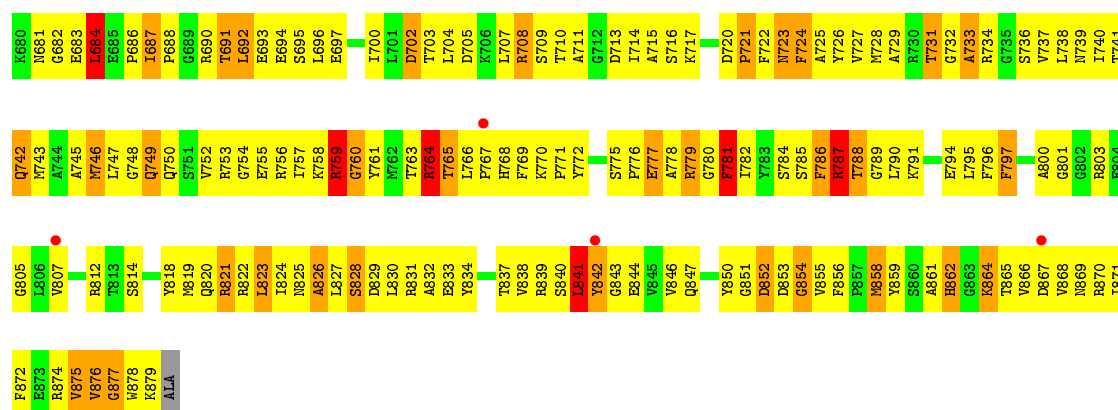




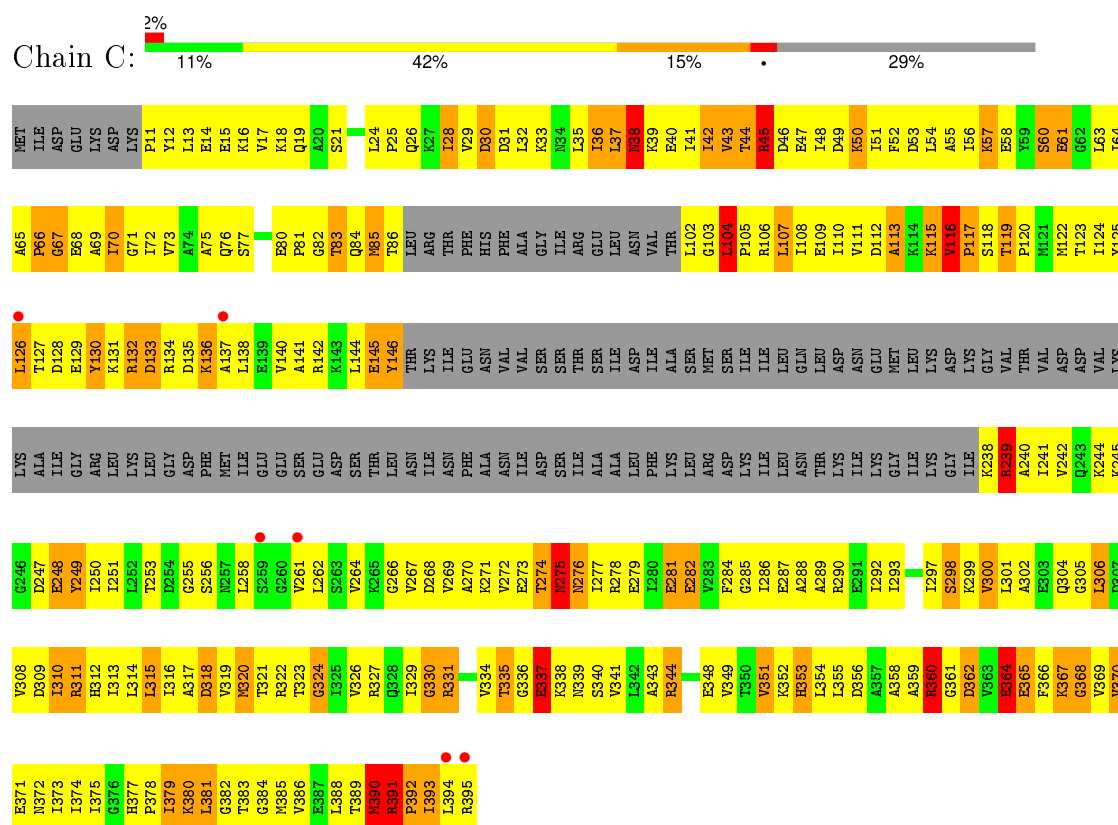


### • Molecule 1: DNA-directed RNA polymerase subunit A

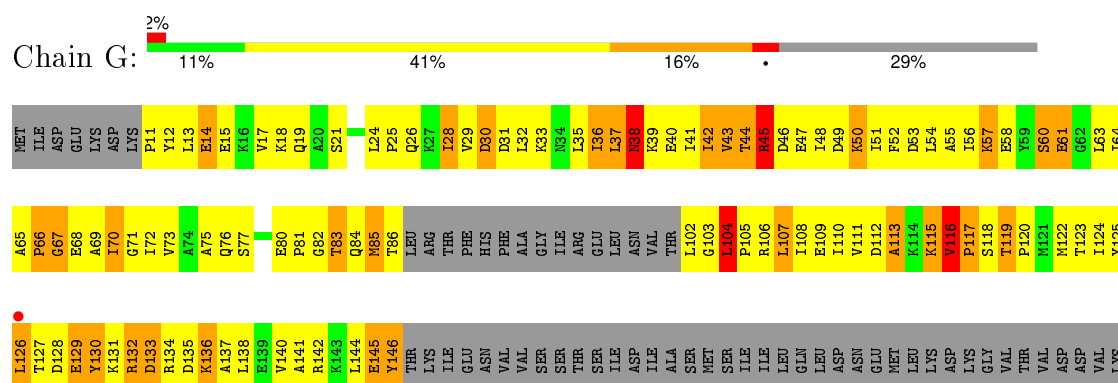


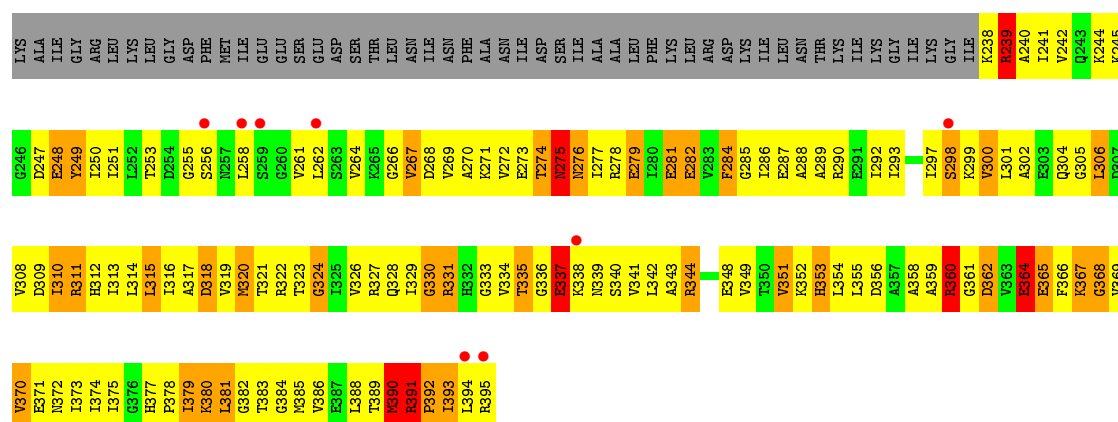


• Molecule 2: DNA-directed RNA polymerase subunit A''

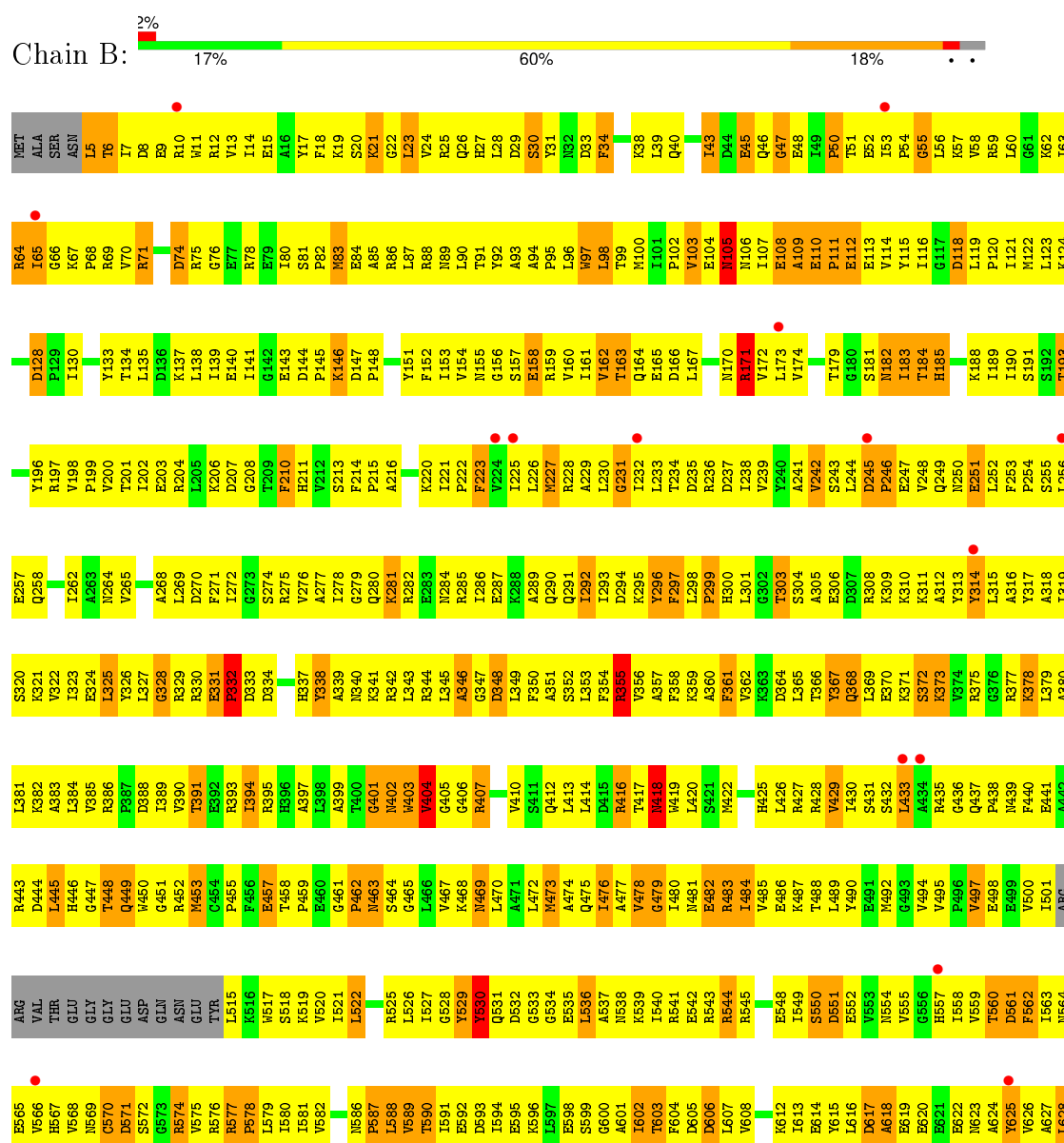


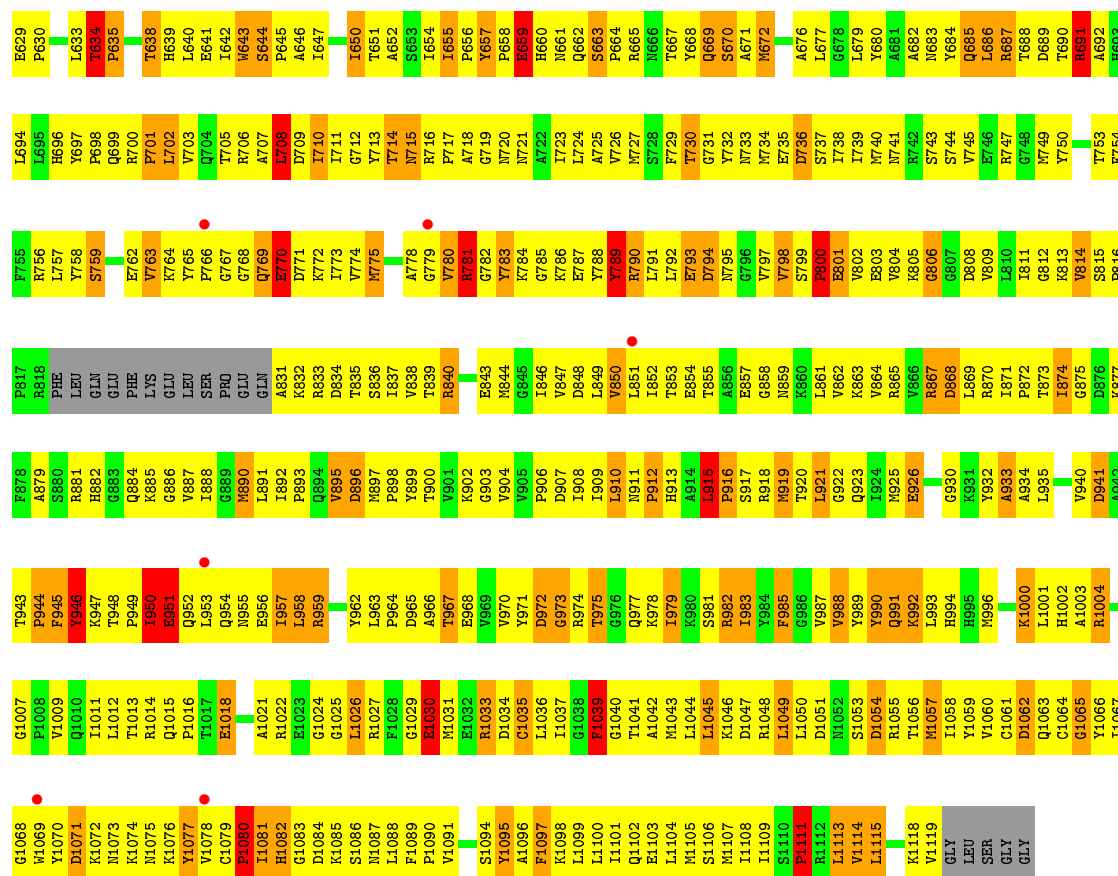
• Molecule 2: DNA-directed RNA polymerase subunit A''



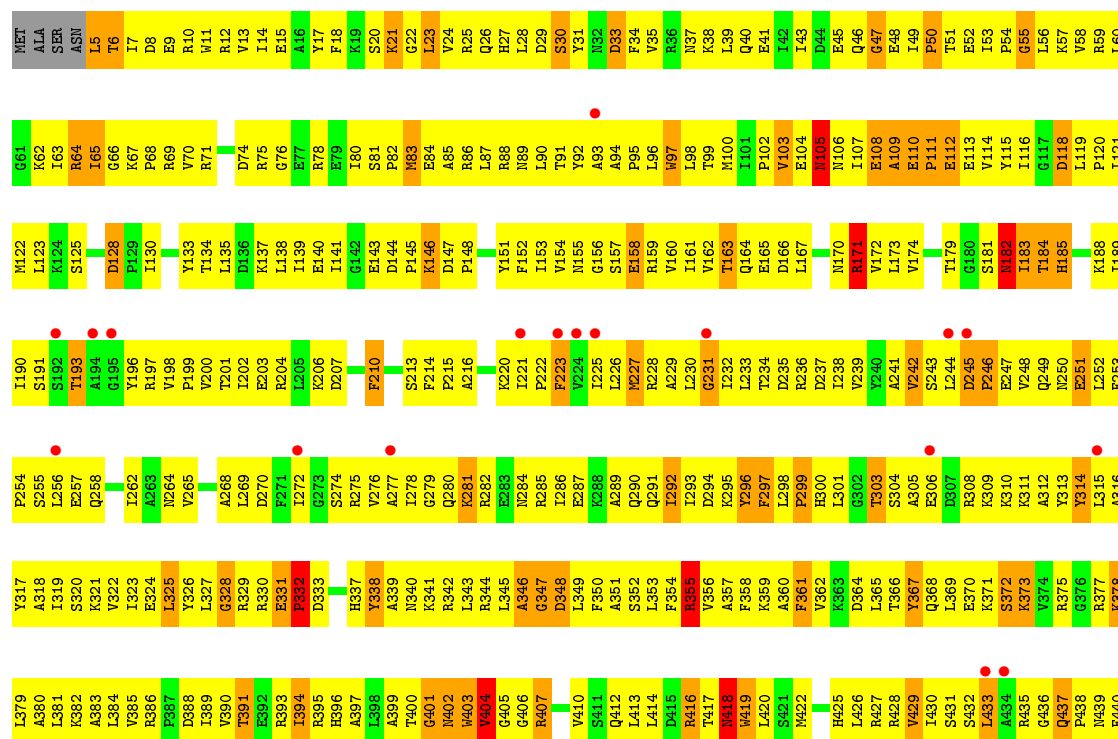


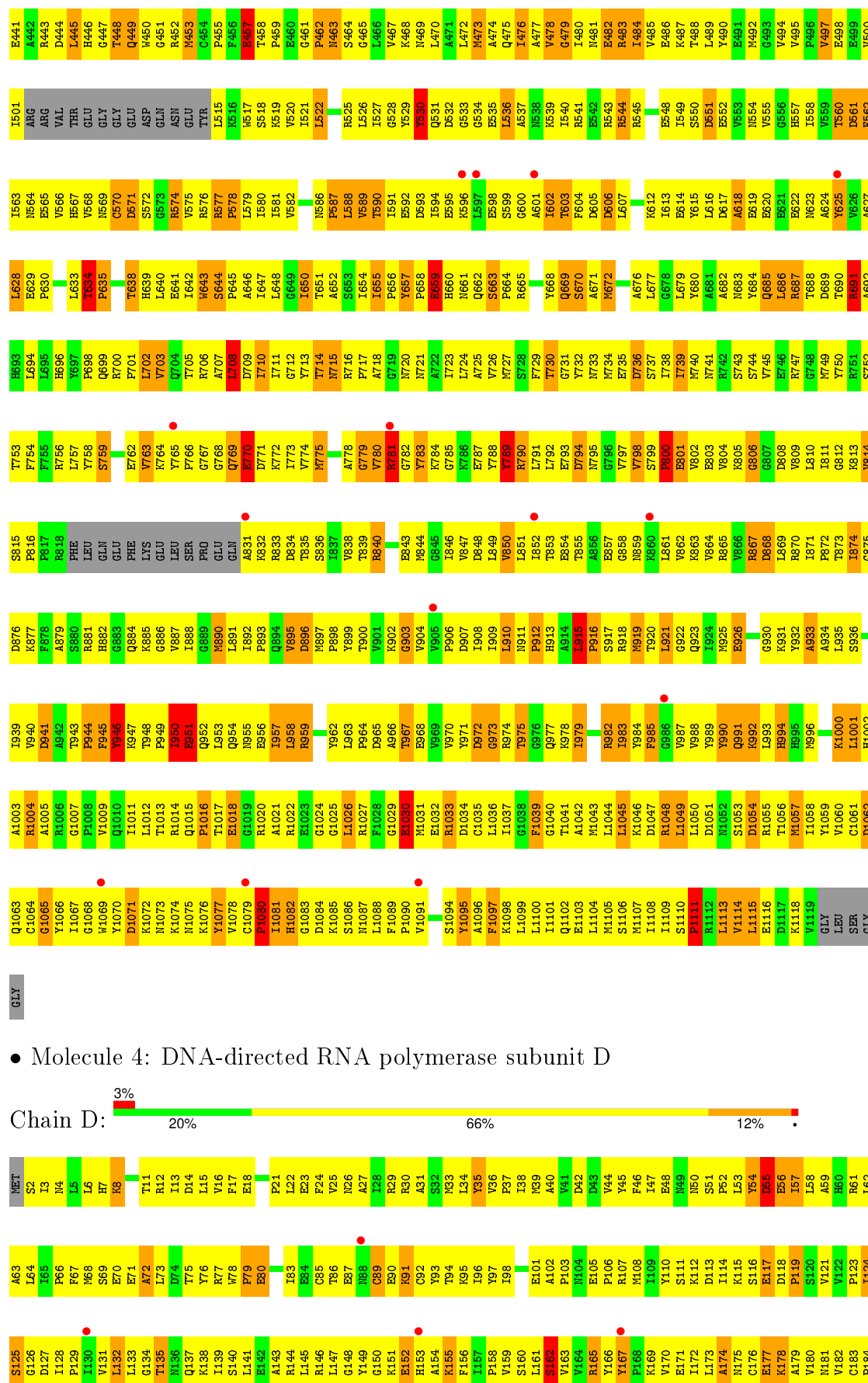
### • Molecule 3: DNA-directed RNA polymerase subunit B

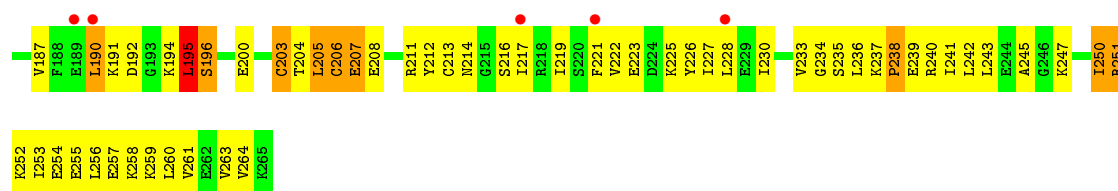




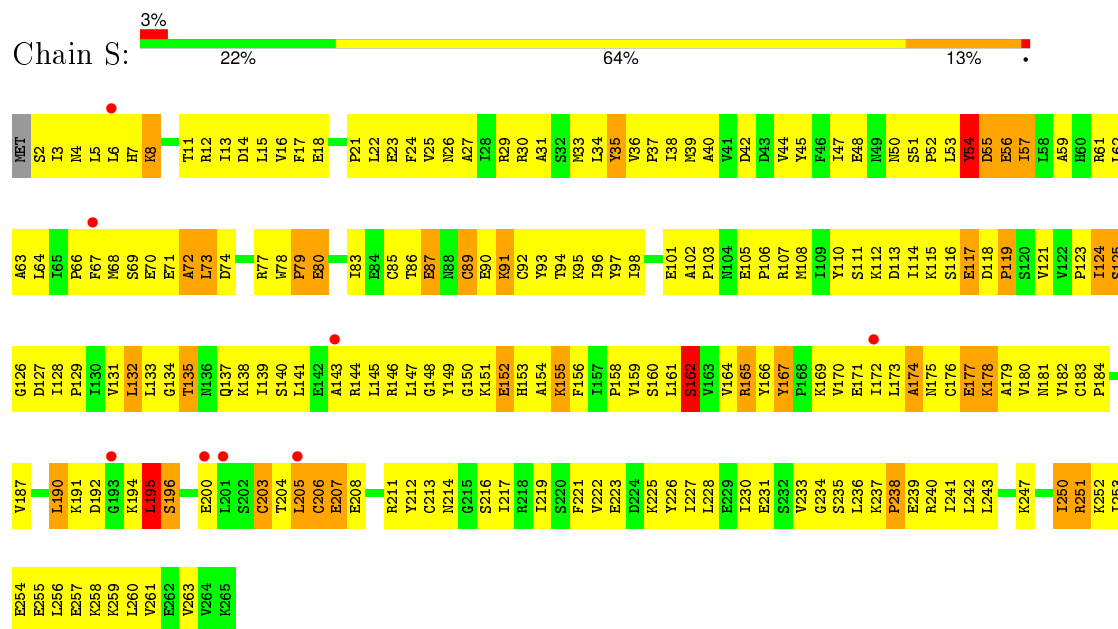
### • Molecule 3: DNA-directed RNA polymerase subunit B



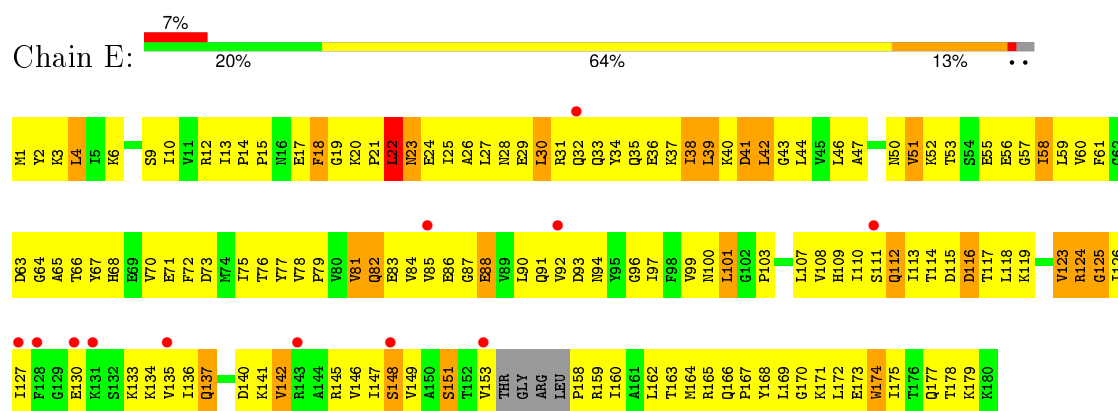




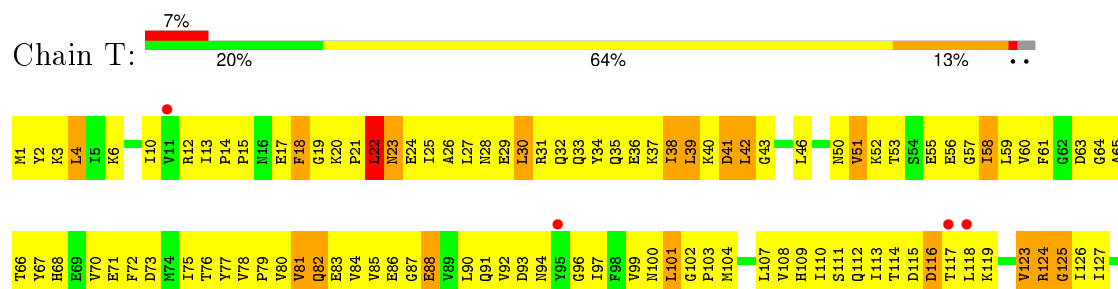
• Molecule 4: DNA-directed RNA polymerase subunit D

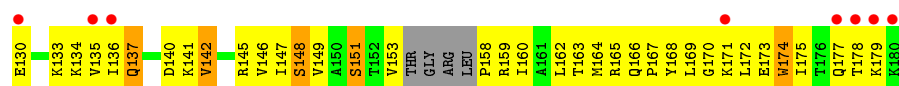


• Molecule 5: DNA-directed RNA polymerase subunit E

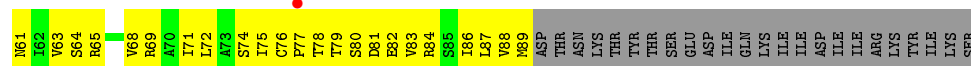
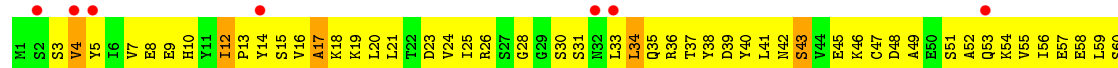
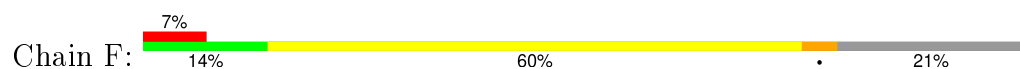


• Molecule 5: DNA-directed RNA polymerase subunit E

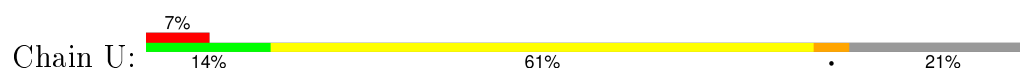




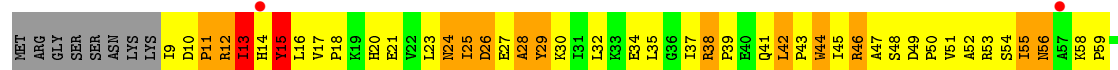
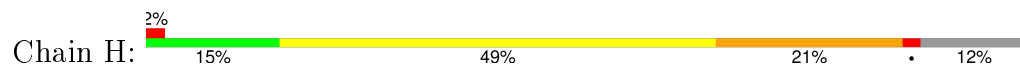
• Molecule 6: DNA-directed RNA polymerase subunit F



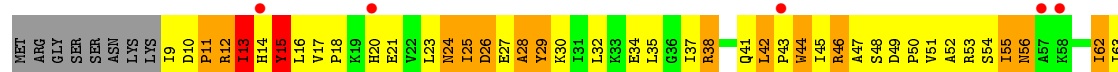
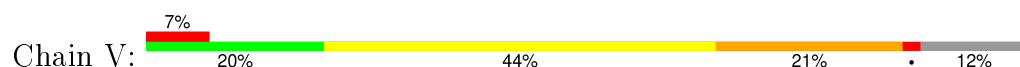
• Molecule 6: DNA-directed RNA polymerase subunit F



• Molecule 7: DNA-directed RNA polymerase subunit H

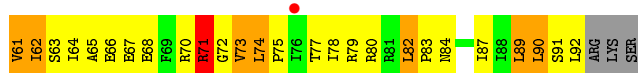


• Molecule 7: DNA-directed RNA polymerase subunit H



• Molecule 8: DNA-directed RNA polymerase subunit K

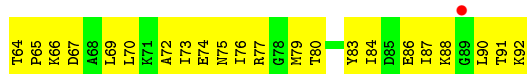
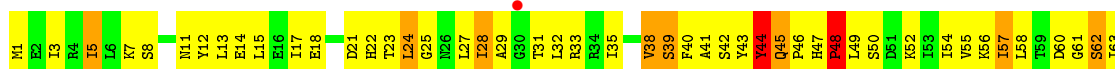




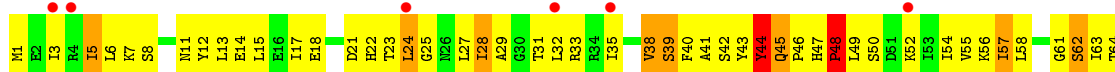
• Molecule 8: DNA-directed RNA polymerase subunit K



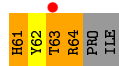
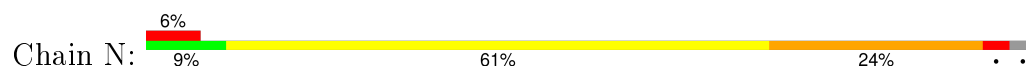
• Molecule 9: DNA-directed RNA polymerase subunit L



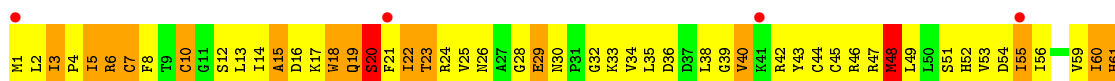
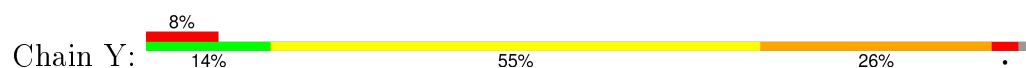
• Molecule 9: DNA-directed RNA polymerase subunit L



• Molecule 10: DNA-directed RNA polymerase subunit N



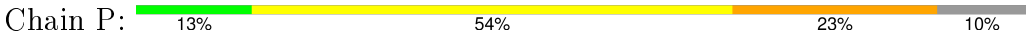
• Molecule 10: DNA-directed RNA polymerase subunit N



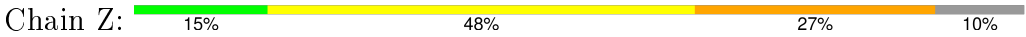




● Molecule 11: DNA-directed RNA polymerase subunit P



● Molecule 11: DNA-directed RNA polymerase subunit P



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.82Å 201.24Å 196.05Å 90.00° 100.92° 90.00°	Depositor
Resolution (Å)	39.79 – 3.40 39.79 – 3.40	Depositor EDS
% Data completeness (in resolution range)	80.4 (39.79-3.40) 80.3 (39.79-3.40)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.274 , 0.343 0.271 , 0.275	Depositor DCC
$R_{free}$ test set	5323 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	79.3	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 105618 reflections	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	48122	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, F3S, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.48	0/6306	0.80	4/8539 (0.0%)
1	Q	0.46	0/6306	0.79	4/8539 (0.0%)
2	C	0.46	0/2189	0.81	0/2947
2	G	0.43	0/2189	0.80	0/2947
3	B	0.46	0/8810	0.79	5/11921 (0.0%)
3	R	0.45	0/8810	0.79	3/11921 (0.0%)
4	D	0.40	0/2152	0.68	0/2911
4	S	0.37	0/2152	0.67	0/2911
5	E	0.38	0/1423	0.69	0/1919
5	T	0.37	0/1423	0.69	0/1919
6	F	0.35	0/701	0.63	0/949
6	U	0.35	0/701	0.62	0/949
7	H	0.44	0/625	0.76	0/848
7	V	0.41	0/625	0.76	0/848
8	K	0.50	0/667	0.82	0/903
8	W	0.49	0/667	0.81	0/903
9	L	0.39	0/733	0.72	0/986
9	X	0.38	0/733	0.72	0/986
10	N	0.38	0/523	0.75	0/705
10	Y	0.37	0/523	0.74	0/705
11	P	0.45	0/354	0.68	0/475
11	Z	0.46	0/354	0.67	0/475
All	All	0.44	0/48966	0.77	16/66206 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	B	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
3	R	0	1
4	D	0	1
4	S	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	841	LEU	CA-CB-CG	7.67	132.93	115.30
1	Q	841	LEU	CA-CB-CG	7.56	132.69	115.30
3	B	436	GLY	N-CA-C	-6.16	97.70	113.10
3	R	436	GLY	N-CA-C	-5.97	98.17	113.10
1	A	508	LEU	N-CA-C	-5.89	95.11	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	314	TYR	Sidechain
4	D	54	TYR	Sidechain
3	R	314	TYR	Sidechain
4	S	54	TYR	Sidechain

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6173	0	6243	1147	0
1	Q	6173	0	6243	1128	0
2	C	2169	0	2288	501	0
2	G	2169	0	2288	526	0
3	B	8645	0	8782	1656	0
3	R	8645	0	8780	1698	0
4	D	2114	0	2145	357	0
4	S	2114	0	2145	348	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	1402	0	1467	222	0
5	T	1402	0	1467	246	0
6	F	694	0	705	129	0
6	U	694	0	705	139	0
7	H	611	0	641	117	0
7	V	611	0	641	125	0
8	K	658	0	692	161	0
8	W	658	0	692	174	0
9	L	723	0	749	94	0
9	X	723	0	749	91	0
10	N	514	0	528	159	0
10	Y	514	0	529	151	0
11	P	346	0	376	63	0
11	Z	346	0	375	58	0
12	A	1	0	0	0	0
12	B	1	0	0	0	0
12	N	1	0	0	0	0
12	P	1	0	0	0	0
12	Q	1	0	0	0	0
12	R	1	0	0	0	0
12	Y	1	0	0	0	0
12	Z	1	0	0	0	0
13	A	1	0	0	0	0
13	Q	1	0	0	0	0
14	D	7	0	0	4	0
14	S	7	0	0	3	0
All	All	48122	0	49230	8272	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 85.

The worst 5 of 8272 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:LYS:NZ	1:A:297:THR:HB	1.42	1.31
1:Q:238:LYS:NZ	1:Q:297:THR:HB	1.43	1.31
1:A:803:ARG:HG2	3:B:444:ASP:HA	1.20	1.17
1:A:308:ARG:HH21	3:B:1099:LEU:HD13	1.10	1.16
3:R:329:ARG:HD2	3:R:562:PHE:HB3	1.24	1.15

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	768/880 (87%)	513 (67%)	136 (18%)	119 (16%)	0	2
1	Q	768/880 (87%)	509 (66%)	141 (18%)	118 (15%)	0	2
2	C	273/392 (70%)	158 (58%)	66 (24%)	49 (18%)	0	2
2	G	273/392 (70%)	161 (59%)	61 (22%)	51 (19%)	0	1
3	B	1084/1124 (96%)	698 (64%)	238 (22%)	148 (14%)	0	3
3	R	1084/1124 (96%)	698 (64%)	237 (22%)	149 (14%)	0	3
4	D	262/265 (99%)	166 (63%)	69 (26%)	27 (10%)	1	7
4	S	262/265 (99%)	167 (64%)	66 (25%)	29 (11%)	0	6
5	E	172/180 (96%)	123 (72%)	31 (18%)	18 (10%)	1	7
5	T	172/180 (96%)	122 (71%)	32 (19%)	18 (10%)	1	7
6	F	87/113 (77%)	56 (64%)	22 (25%)	9 (10%)	1	7
6	U	87/113 (77%)	56 (64%)	23 (26%)	8 (9%)	1	9
7	H	72/84 (86%)	46 (64%)	13 (18%)	13 (18%)	0	1
7	V	72/84 (86%)	44 (61%)	15 (21%)	13 (18%)	0	1
8	K	80/95 (84%)	44 (55%)	19 (24%)	17 (21%)	0	1
8	W	80/95 (84%)	44 (55%)	20 (25%)	16 (20%)	0	1
9	L	90/92 (98%)	64 (71%)	19 (21%)	7 (8%)	1	12
9	X	90/92 (98%)	66 (73%)	17 (19%)	7 (8%)	1	12
10	N	62/66 (94%)	30 (48%)	18 (29%)	14 (23%)	0	1
10	Y	62/66 (94%)	31 (50%)	18 (29%)	13 (21%)	0	1
11	P	41/48 (85%)	24 (58%)	10 (24%)	7 (17%)	0	2
11	Z	41/48 (85%)	23 (56%)	11 (27%)	7 (17%)	0	2
All	All	5982/6678 (90%)	3843 (64%)	1282 (21%)	857 (14%)	0	3

5 of 857 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	56	GLN
1	A	58	CYS
1	A	64	THR
1	A	65	LEU
1	A	194	ILE

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	675/766 (88%)	580 (86%)	95 (14%)	4	22
1	Q	675/766 (88%)	583 (86%)	92 (14%)	5	24
2	C	237/338 (70%)	201 (85%)	36 (15%)	3	19
2	G	237/338 (70%)	199 (84%)	38 (16%)	3	16
3	B	937/965 (97%)	807 (86%)	130 (14%)	4	23
3	R	937/965 (97%)	810 (86%)	127 (14%)	5	24
4	D	241/242 (100%)	224 (93%)	17 (7%)	18	58
4	S	241/242 (100%)	223 (92%)	18 (8%)	17	54
5	E	156/159 (98%)	142 (91%)	14 (9%)	12	45
5	T	156/159 (98%)	142 (91%)	14 (9%)	12	45
6	F	82/106 (77%)	79 (96%)	3 (4%)	41	77
6	U	82/106 (77%)	79 (96%)	3 (4%)	41	77
7	H	67/75 (89%)	54 (81%)	13 (19%)	2	8
7	V	67/75 (89%)	55 (82%)	12 (18%)	2	11
8	K	72/84 (86%)	57 (79%)	15 (21%)	1	6
8	W	72/84 (86%)	57 (79%)	15 (21%)	1	6
9	L	81/81 (100%)	75 (93%)	6 (7%)	17	55
9	X	81/81 (100%)	75 (93%)	6 (7%)	17	55
10	N	58/60 (97%)	50 (86%)	8 (14%)	4	23
10	Y	58/60 (97%)	49 (84%)	9 (16%)	3	18

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	P	39/43 (91%)	31 (80%)	8 (20%)	1	7
11	Z	39/43 (91%)	31 (80%)	8 (20%)	1	7
All	All	5290/5838 (91%)	4603 (87%)	687 (13%)	5	25

5 of 687 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	K	23	TRP
1	Q	425	LEU
6	U	5	TYR
8	K	74	LEU
1	Q	52	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 175 such sidechains are listed below:

Mol	Chain	Res	Type
7	H	41	GLN
1	Q	376	ASN
5	T	109	HIS
8	K	42	GLN
1	Q	56	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 10 are monoatomic - leaving 2 for Mogul analysis.



In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
14	F3S	D	1001	4	0,9,9	0.00	-	0,15,15	0.00	-
14	F3S	S	1001	4	0,9,9	0.00	-	0,15,15	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
14	F3S	D	1001	4	-	0/0/24/24	0/0/3/3
14	F3S	S	1001	4	-	0/0/24/24	0/0/3/3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	D	1001	F3S	4	0
14	S	1001	F3S	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	776/880 (88%)	0.03	12 (1%) 76 71	23, 75, 132, 186	0
1	Q	776/880 (88%)	0.14	22 (2%) 56 52	29, 87, 145, 202	0
2	C	279/392 (71%)	0.06	6 (2%) 65 60	29, 81, 153, 190	0
2	G	279/392 (71%)	0.08	9 (3%) 51 47	42, 97, 153, 181	0
3	B	1090/1124 (96%)	0.07	21 (1%) 70 64	24, 78, 145, 196	0
3	R	1090/1124 (96%)	0.14	32 (2%) 55 50	36, 84, 149, 196	0
4	D	264/265 (99%)	0.14	9 (3%) 49 44	44, 94, 144, 179	0
4	S	264/265 (99%)	0.16	8 (3%) 54 49	61, 111, 157, 192	0
5	E	176/180 (97%)	0.43	12 (6%) 20 19	39, 112, 189, 202	0
5	T	176/180 (97%)	0.39	12 (6%) 20 19	57, 113, 176, 202	0
6	F	89/113 (78%)	0.37	8 (8%) 12 11	73, 142, 182, 196	0
6	U	89/113 (78%)	0.49	8 (8%) 12 11	93, 141, 184, 201	0
7	H	74/84 (88%)	0.23	2 (2%) 58 53	46, 90, 137, 165	0
7	V	74/84 (88%)	0.24	6 (8%) 15 14	70, 101, 159, 198	0
8	K	82/95 (86%)	0.03	2 (2%) 62 57	30, 72, 118, 154	0
8	W	82/95 (86%)	0.14	1 (1%) 81 75	47, 82, 139, 189	0
9	L	92/92 (100%)	0.17	2 (2%) 65 60	42, 81, 125, 200	0
9	X	92/92 (100%)	0.31	9 (9%) 10 10	52, 102, 140, 159	0
10	N	64/66 (96%)	0.09	4 (6%) 23 22	60, 90, 124, 174	0
10	Y	64/66 (96%)	0.17	5 (7%) 16 15	61, 98, 145, 185	0
11	P	43/48 (89%)	0.09	0 100 100	48, 101, 137, 155	0
11	Z	43/48 (89%)	-0.02	0 100 100	64, 100, 144, 180	0
All	All	6058/6678 (90%)	0.13	190 (3%) 52 48	23, 88, 153, 202	0

The worst 5 of 190 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	434	ALA	9.9
6	U	33	LEU	7.3
1	Q	212	LEU	6.2
6	U	89	MET	6.2
6	U	2	SER	6.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
12	ZN	N	1001	1/1	0.99	0.17	1.26	93,93,93,93	0
12	ZN	Z	1001	1/1	1.00	0.23	0.14	106,106,106,106	0
12	ZN	B	2001	1/1	0.98	0.18	0.12	91,91,91,91	0
12	ZN	P	1001	1/1	0.99	0.22	-0.10	103,103,103,103	0
12	ZN	Y	1001	1/1	0.99	0.16	-0.31	93,93,93,93	0
14	F3S	S	1001	7/7	0.99	0.19	-0.66	111,111,112,113	0
14	F3S	D	1001	7/7	0.98	0.19	-0.80	79,80,80,80	0
12	ZN	A	1002	1/1	0.99	0.14	-1.08	87,87,87,87	0
12	ZN	R	2001	1/1	0.98	0.11	-1.34	101,101,101,101	0
12	ZN	Q	1002	1/1	0.99	0.11	-1.49	89,89,89,89	0
13	MG	A	1003	1/1	0.89	0.23	-	57,57,57,57	0
13	MG	Q	1003	1/1	0.98	0.29	-	60,60,60,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.